(12) UK Patent Application (19) GB (11) 2 341 059 (13) A

(43) Date of A Publication 01.03.2000

(21) Application No 9821994.2

(22) Date of Filing 08.10.1998

(30) Priority Data

(31)	9818873	(32)	28.08.1998	(33)	GB
(31)	9819175	(32)	01.09.1998		

(71) Applicant(s)
Nokia Oy
(Incorporated in Finland)
Keilalahdentie 4, 12150 Espoo, Finland

(72) Inventor(s)
Tom Soderlund
Jukka Immonen

(74) Agent and/or Address for Service
Nokia IPR Department
Nokia House, Summit Avenue, Southwood,
FARNBOROUGH, Hampshire, GU14 0NG,
United Kingdom

(51) INT CL^7
H04L 12/56 12/26 29/06

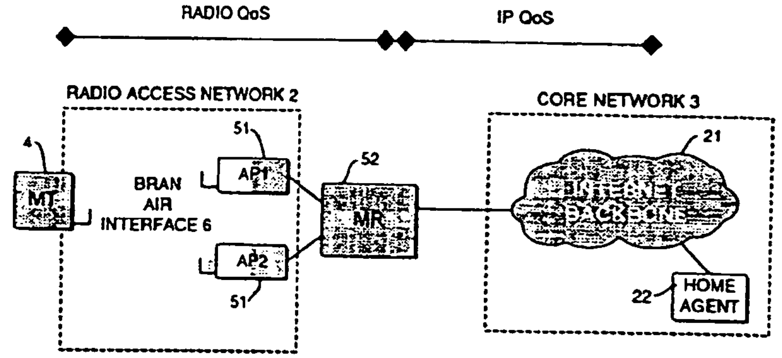
(52) UK CL (Edition R)
H4P PEUX PPS

(56) Documents Cited
GB 2248535 A EP 0848527 A1

(58) Field of Search
UK CL (Edition Q) H4P PEUX PFD PPG PPS
INT CL^6 H04L 12/24 12/26 12/56 29/06
Online : EPODOC

(54) Abstract Title
Internet protocol flow detection

(57) An IP flow detector (51; 52) is provided which supports differentiated services in an IP network, such as a wireless IP network. In one embodiment, the detector is arranged to detect a flow type which can be identified by fields in the basic IPv6 header and an extension header. Likewise a method of detecting such IP flows is provided. If there is no "flow label" availability, then TCP/VDP part information is used. Lacking the latter, some other field in a lower layer header indicative of packet management criteria is used, eg. a security parameter index of the encapsulating security payload header where IP encryption is used. Failing this, flows are identified by source and destination IP addresses.

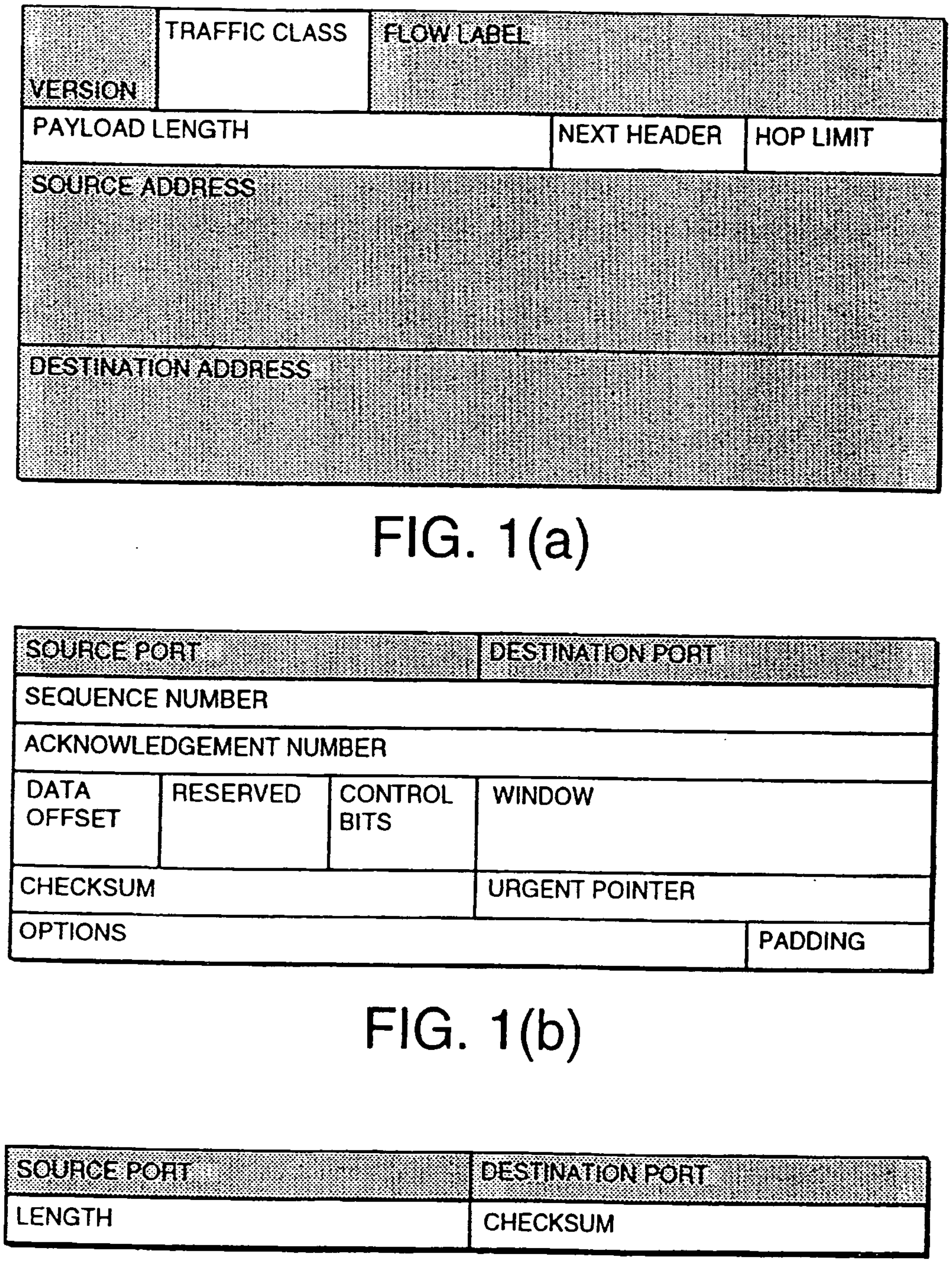


MT = MOBILE TERMINAL
AP = ACCESS POINT
MR = MOBILITY ENHANCED IP ROUTER

FIG. 3

GB 2 341 059 A

At least one drawing originally filed was informal and the print reproduced here is taken from a later filed formal copy.



VERSION	TRAFFIC CLASS	FLOW LABEL

PAYLOAD LENGTH	NEXT HEADER	HOP LIMIT

SOURCE ADDRESS
DESTINATION ADDRESS

FIG. 1(a)

SOURCE PORT	DESTINATION PORT

SEQUENCE NUMBER
ACKNOWLEDGEMENT NUMBER

DATA OFFSET	RESERVED	CONTROL BITS	WINDOW

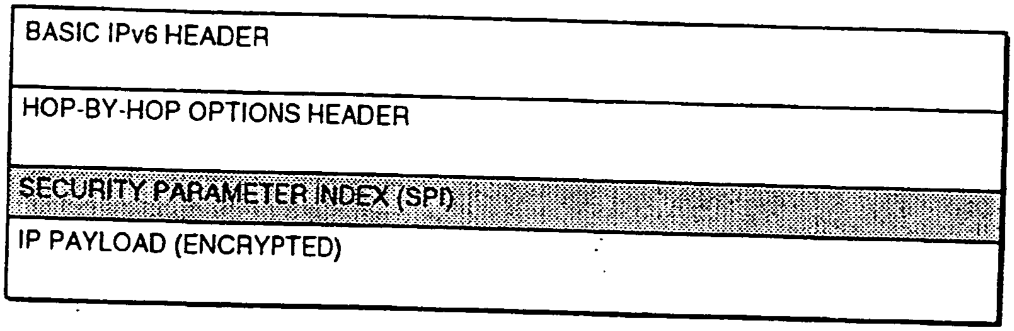
CHECKSUM	URGENT POINTER
OPTIONS	PADDING

FIG. 1(b)

SOURCE PORT	DESTINATION PORT
LENGTH	CHECKSUM

FIG. 1(c)

FIG. 1, IPv6, TCP AND UDP HEADER FORMATS



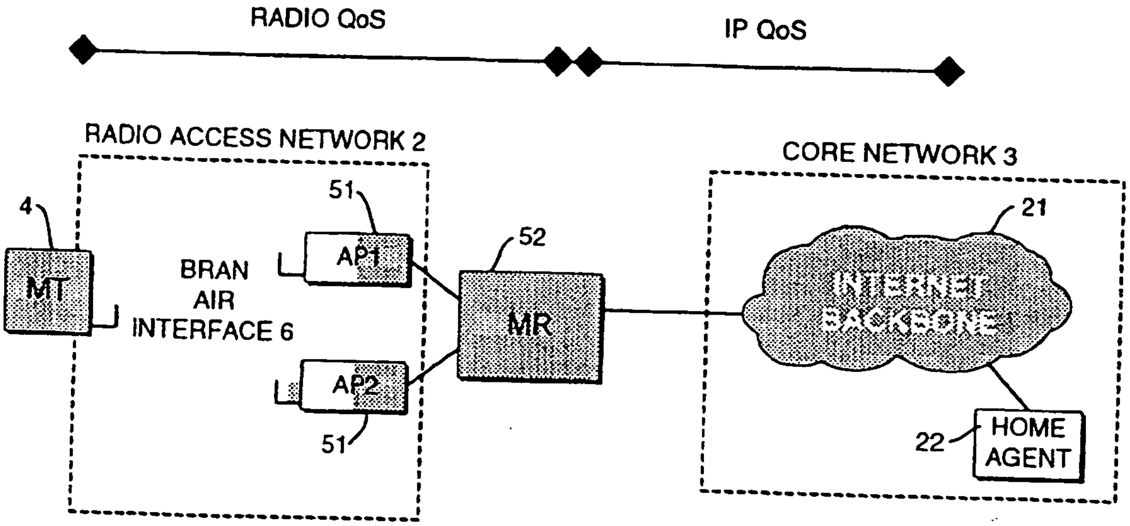
BASIC IPv6 HEADER

HOP-BY-HOP OPTIONS HEADER

SECURITY PARAMETER INDEX (SPI)

IP PAYLOAD (ENCRYPTED)

FIG. 2 IPv6 HEADER FORMAT WHEN ESP USED



MT = MOBILE TERMINAL
AP = ACCESS POINT
MR = MOBILITY ENHANCED IP ROUTER

FIG. 3

2341059

Internet protocol flow detection

The present invention relates to a mechanism for differentiating between internet protocol (IP) packets, namely IP flow detection. In particular, it relates to a method and apparatus which provide flow detection in packet data transmission.

Such a mechanism is used, for example, in wireless internet protocol (IP) networks.

The term "Internet" is commonly used to describe an information resource from which information can be retrieved from a data processor, such as a personal computer (PC). The data processor communicates with the other nodes in the network via e.g., a modem hooked to a telecommunication network. The data processor may be connected to the network also by other means like a direct data network connection. This information resource is worldwide. The Internet is made operable by defining certain data communication standards and protocols, such as TCP (transfer control protocol), UPD (user datagram protocol), and IP (Internet protocol), which are used for controlling data transmission between numerous parts of the Internet. The TCP is involved with providing communicating processes means to transfer a stream of data reliably. The UDP is involved with providing communicating processes means to transfer datagrams reliably. The IP is the common ground for communicating between the nodes in the IP network. The lower part of IP can be modified to suit specific network environments while the upper part of IP as well as the TCP and the UDP above it remain the same everywhere. This way a global uniform network is possible although it has been built on local networks that are based on different technologies. The currently used versions of the Internet protocol are IPv4 and IPv6. IPv4 is defined in RFC791 and IPv6 is defined in the Ipv6 specification dated 8th June 1998.

Thanks to the growing popularity of open data systems, the Transmission Control Protocol/Internet Protocol (TCP/IP) communication protocol has become a generally used protocol whereby computers of different sizes and brands can communicate with each other. TCP/IP support is currently available for almost any operating system and almost any local network environment. The network layer protocol of TCP/IP, the Internet Protocol IP, is intended to be routed by gateways, *i.e.* routers. The routing is conducted by means of IP addresses and routing tables. Thanks to the TCP/IP protocol suite, the applications running in the hosts connected to the Internet are able to communicate even though the hosts were located in different continents or even space.

The rapid evolution of the Internet services has created a strong need for broadband networks with high data rate and Quality of Service (QoS). Video broadcasting and other multimedia distribution services are evolving rapidly. The users are willing to access these services also in the wireless environment. Currently, in the fixed IP network, IP packets are typically sent as best effort data traffic. In the case of network congestion, all data streams are handled with equal priority which may have a dramatic effect on multimedia services. Two main problems exist: firstly the current wireless networks do not provide sufficient QoS mechanism, and secondly, the existing wireless networks are not capable of serving several simultaneous connections with high data rate and QoS requirements. To meet the increased customer requirements, new wireless broadband network techniques are required.

The Internet Engineering Task Force (IETF) is an organisation involved with the development of the architecture, protocols and operation of the Internet. IEFT has defined two different QoS concepts: integrated and differentiated services, for providing a standard mechanism for supporting real time applications in IP networks. Integrated services is based on an abstract flow model with reservation protocol (RSVP, RFC2205) and admission control. The network reserves statically resources for real time connections in each network device, and hence is not optimally efficient. Consequently the differentiated services concept was developed. This concept is based on the use of an IP header for indicating the requested service class (called per-hop behaviour) for the packet. As a result, each IP packet carries QoS information and no reservations are required. Whilst the IETF suggest the use of an IP header for indicating the QoS, the actual packet handling mechanisms will not be standardised.

One mechanism used for differentiating between IP packets is IP flow detection. The flow detection concept comprises the monitoring of IP traffic to be able to detect packets flowing frequently between two communicating processes (IP applications). Such packets establish an IP flow.

The flow detection entity (called the flow classifier) monitors the IP packets and some specific IP header fields in detecting the flows. There are several header fields (parameters) which can be used in flow detection. Figures 1(a) to (c) present the IPv6, TCP and UDP protocol headers respectively and the header fields applicable to flow detection.

As can be seen from Figure 1(a), the IPv6 header block consists of the following elements:

Version	IP version of 4 bits (=6)

Traffic Class 8 bit priority,

Flow label 20 bit label for identifying the connection in the application layer,

Payload length 16 bit integer indicating the length of the payload, *i.e.* the length of the packet after the header in bytes,

Next header data of 8 bits determining the header immediately following the IPv6 header,

Hop limit integer counter of 8 bits which is reduced by one at the each device (node) which transmits the packet further; the packet is rejected if the value is reduced to zero,

Source address the 128 bit address of the sender of the original packet,

Destination address the 128 bit address of the intended recipient.

The header is followed by the payload block, *i.e.* the actual information to be transmitted.

In this IP version, if a packet is provided with a non-zero IPv6 flow label, then the flow label in the IPv6 header together with the source address is a flow identifier (first flow identification type), and directly distinguishes the different IP sessions. The flow label is used by the applications to "mark" the packets belonging to their IP flow. However, this flow label is not available in all IP headers (e.g. IPv4). Moreover, if a system does not support it, the flow label is set to zero. In these instances, fields of an upper layer protocol header can be used as flow identifiers instead, to distinguish different IP sessions. For example, as is shown in Figure 1b, an alternative flow ID can include TCP/UDP source and destination port information to distinguish different IP sessions together with source address and destination address from the IP header (second flow identification type).

The present inventors have realised the need for a method of flow detection which provides an improved flexibility (for example so its use is not solely dependent upon the use of current protocols).

According to one aspect of the present invention there is provided a method for detecting an IP flow in flow label deprived packet data transmission, comprising monitoring a set of fields in a lower layer header of the packets to detect an IP flow, wherein monitoring the set of fields comprises monitoring an source address field; monitoring a destination address field; and monitoring a further field indicative of packet management criteria. A lower layer header is generally one from OSI layers 1 to 3 (physical, data link and network layers) and may for example be an IP header or an IP extension header. This IP flow detection method supports the aforementioned QoS concepts for example, when an upper layer header is not available. Consequently, it provides effective flow management.

According to another aspect of the present invention, there is provided a method of detecting an IP flow in packet data transmission, comprising selecting a set of fields to be monitored to detect an IP flow; and monitoring the selected set of fields in a header of the packets to detect an IP flow; wherein the set of fields is selected from a first set comprising a flow label field and a source address field from a lower layer header of the packets; a second set comprising a source address field and a destination address field from a lower layer header of the packets and a source port field and a destination port field from an upper layer header of the packets; and a third set comprising the source address field, destination address field, and a further field indicative of packet management criteria (other than the flow label field) from the lower layer header of the packets. This IP flow method uses different options, depending on what headers are available. Again, these options provide a flexible method of providing effective flow management.

Preferably the set of fields are selected in the aforementioned order, as they are ranked having consideration to the complexity and load in routing devices. That is, the lower the priority the increased complexity and load in the routing devices.

The method may be used for detecting an IP flow encrypted packet data transmission. This is one example when not all the headers are accessible, and in which the present invention is particularly useful. In such a case, the step of monitoring the further field comprises monitoring a security field, such as a security parameter index of the encapsulated security payload header.

According to a further aspect of the present invention, there is provided the use of a set of fields in a lower layer header of data packets as a flow identifier, wherein the set of fields comprises source and destination address fields and a field indicative of packet management criteria (other than a flow label field).

According to yet another aspect of the present invention, there is provided an IP flow detector which implements any of the methods above.

Embodiments of the present invention will now be described, by way of example, of which:

Figure 1a illustrates an IPv6 header format;
Figure 1b illustrates a TCP header format;
Figure 1c illustrates a UDP header format;

Figure 2 illustrates an IPv6 header format when encapsulated security payload (ESP) is used; and

Figure 3 illustrates a general radio system architecture in which the method and apparatus of the present invention may be implemented.

In the present invention, flow detection is made more flexible by the provision of an alternative flow identifier to the first and second types mentioned above, which caters for the situation where a flow label and/or port addresses are not available, and yet still supports differentiated services. A flow label may not be available for the following reasons, for example. Firstly, the system may be using a protocol which does not define a flow label field (e.g. IPv4). Secondly, the system may not support flow labels, despite using a protocol which defines a flow label field (e.g. IPv6). In this latter case, the flow label value is set to zero.

Further, an example of a situation when neither a flow label nor port addresses are available, may be when the data packet is encrypted for security reasons. Encryption of IP packets is discussed in the draft on IP Encapsulating Security Payload (ESP) by Stephen Kent and Randall Atkinson, dated March 1998. Figure 2 of the accompanying drawings illustrates an ESP extension header.

IP encryption utilises Encapsulating Security Payload (ESP), ESP provides means for encrypting the contents of an IP packet. When ESP based encryption is used (transport-mode ESP), the processing nodes can interpret only the basic IP header and the extension headers preceding the ESP header. This is because transport-mode encapsulates a transport-layer (e.g., UDP, TCP or ICMP) frame inside the ESP. In transport-mode ESP, the ESP header follows the end-to-end headers (e.g., Authentication Header) and immediately precedes an upper layer (e.g., UDP, TCP, ICMPv6) header.

The encrypted data includes both the protected ESP header fields and the protected user data, which is either an entire IP packet or an upper-layer protocol frame (e.g., TCP or UDP). Thus, the second flow identification type cannot be used because there is no port information available. If, in addition, the flow label is not set, effective flow management cannot be achieved.

In a preferred embodiment of the present invention, an alternative flow identifier is used which makes use of a source address field, destination address field of the basic IPv6 header and also a further field which is indicative of packet management criteria, and thus provide effective flow detection. On their own, the source and destination addresses do not provide any differentiation and thus do not provide efficient and effective flow management. (Hosts may have several applications running, each application requiring different treatment in the network). In contrast, in the present example of the invention, the Security Parameter Index (SPI) value from the ESP header is used for flow identification. This header is indicative of packet management criteria.

The SPI value together with the source and destination addresses can be used to identify different IP flows. The SPI field is a 32-bit arbitrary value which together with the destination IP address and security protocol (here ESP) identifies the Security Association for the packet. Figure 2 shows the IP packet format when ESP encryption is used (the IPv6 header fields applicable to flow detection are highlighted in Figure 1).

When IP encryption is applied (transport-mode ESP), only the information contained in the basic IPv6 header and the extension headers preceding the ESP header can be used for flow detection. From the basic header, flow label, source address and destination address parameters are usable. If flow label values are not used, flow identification granularity can be improved by using the SPI value from the ESP header, together with the source and destination IP addresses to identify a flow. That is, the information available to differentiate IP flows from each other can be increased.

Thus the solution for detecting flows also for secured IP traffic to provide a third flow identification type as follows:

- type 3: source address + destination address +[SPI]

Instead of using only the source and destination IP addresses to identify the flow the packet belongs to, it is proposed to also use the SPI value, if there is one available in the extension headers. The flow identification granularity depends on the granularity of the respective Security Association that is identified by the SPI value, the destination IP address and the security protocol (ESP). For example, there may exist several Security Associations between two hosts or only one Security Association. If there was only one Security Association between the two hosts all the different TCP flows would be detected by the proposed algorithm as only one flow because the identifying information in the packets of these flows would be the same (same IP addresses and same SPI value in the ESP header). If however there are several Security Associations between these two hosts e.g., per TCP port basis, the proposed flow detection algorithm would detect each TCP flow as a separate flow because the packets of each flow would have different identifying information (same IP addresses, different SPI values in the ESP header).

The function that handles the flow detection is called the flow classifier. The flow classier decides when subsequent IP packets with the same flow identifying information should be considered as a flow. Usually some measurements are made at this point to ensure that the packets that belong to the detected flow will be handled specially. For example, in the case of a wireless media, the flow could be bound to a radio connection.

The flow classifier decides also when the flow should be terminated. The flow will be terminated when no packets with the same flow identifying information are detected within a reasonable time. At this point the resources will be freed that were reserved for the flow when it was detected. For example, in the case of a wireless media, the radio connection would be freed that was reserved for the packets that belonged to the terminated flow.

Flow classification works so that the IP traffic and specific header fields are monitored in order to detect new flows. In a preferred embodiment of the present invention, the classifier specifies a number of different flow types which can be used, depending on the IP and transport protocol header fields. For example, the following four different flow types can be specified:

1. Flows identified by flow labels (type 1)
2. Flows identified by TCP/UDP port numbers (type 2)
3. Flows identified by the source and destination IP addresses + the security parameter index (type 3)
4. Flows identified by the source and destination IP addresses (type 4)

The first option can be applied if the applications are able to use the IPv6 flow label to mark the different IP sessions. If such advanced applications are not available, and if TCP/UDP port information is available, flow type 2 is selected. In case IP encryption is used, the second option cannot be applied since the port information is encrypted. In such a case, security parameter index (SPI) is used with source and destination addresses to identify possible flows. If no TCP/UDP port information, flow labels or SPI parameters are available, the only option is to look just for the source and destination IP addresses and separate flows between hosts (the first two options separate flows in the granularity of communicating processes).

Each flow type specifies the set of fields from the IP packet header that are used to identify a flow. The set of the header fields identifying a particular flow is called the flow identifier. Depending on the flow type, the flow identifiers contain the following fields:

- Type 1: source address + flow label
- Type 2: source address + destination address + protocol (next header) + source port + destination port
- Type 3: source address + destination address + security parameter index (SPI)
- Type 4: source address + destination address

In this embodiment, the flow types have been prioritised in the above order to minimise the required load and processing. The first case can be applied when the packets have a nonzero IPv6 flow label, distinguishing directly the different IP sessions. This uses the least processing. As mentioned above, if the flow label is not available but TCP/UDP port information is available instead, the second case is selected. However, this second option requires UDP/TCP header processing allowing efficient flow management but at the same time increasing the complexity and load in routing devices. If neither flow label nor TCP/UDP port information is available, the flows can be identified by the source and destination addresses and the SPI. Since this SPI value is in an extension header, it again requires an increased load over type 1.

A flow detector can differentiate between these four flow types, and based on the flow classification mechanism bind each flow type to a flow. Three different flow classifier mechanisms which may be applied in the present system are:

- X/Y classifier, meaning X packets (with the same flow identifier) in Y seconds resulting in a new flow
- Protocol classifier which simply assigns all TCP packets to flows
- Port classifier, using transport layer port numbers to decide which flows to bind.

The X/Y classifier is the preferred choice as it is the only one which supports flow types 1 and 2.

Typical flow detection criteria for the X/Y classifier are listed in table 1 below. The table gives values for X and Y in a function of different amount of flow space available (e.g. the flow space may refer to the amount of radio connections required in a wireless internet system). Expected performance means the portion of packets switched to flows.

As can be seen, the values are somewhat different in different environments. Therefore, it should be possible to change easily the values of X and Y in the WFMP implementation.

Table 1: X/Y classifier recommendations

Flow space req.	Gateway	Campus/Enterprise Backbone
1K	Classifier: X = 5 / Y = 15 sec. Flow deletion delay: 30-120 sec. Expected performance: 85%	Classifier: X = 40 / Y = 40 sec. Flow deletion delay: 30-60 sec. Expected performance: 79%
2K	Classifier: X = 5 / Y = 60 sec. Flow deletion delay: 30-120 sec. Expected performance: 90%	Classifier: X = 10 / Y = 45 sec. Flow deletion delay: 30-60 sec. Expected performance: 89%
8K	Classifier: X = 2 / Y = 60 sec. Flow deletion delay: 30-120 sec. Expected performance: 93%	Classifier: X = 5 / Y = 60 sec. Flow deletion delay: 30-60 sec. Expected performance: 92%
16K	Classifier: X = 2 / Y = 60 sec. Flow deletion delay: 30-120 sec. Expected performance: 93%	Classifier: X = 2 / Y = 60 sec. Flow deletion delay: 30-60 sec. Expected performance: 95%
32K	Classifier: X = 2 / Y = 60 sec. Flow deletion delay: 30-120 sec. Expected performance: 93%	Classifier: X = 2 / Y = 60 sec. Flow deletion delay: 30-60 sec. Expected performance: 95%
∞	Classifier: all packets Flow deletion delay: ∞ Expected performance: 99%	Classifier: all packets Flow deletion delay: ∞ Expected performance: 98%

Since the establishment of a TCP connection always contains at least three packets used, and since the flow detection should be based on actual data packets, a minimum value of six for X is considered appropriate (third data packet triggering the flow detection). The value for Y could be 30 seconds.

A flow is deleted after some constant number of seconds of inactivity. When flow classifier detects a new flow, it starts the flow inactivity timer. This timer is re-started each time a packet belonging to that flow is received. Once the timer expires, the flow identifier is removed from the list of monitored packets. Finally, the IP flow is released.

One implementation of the internet protocol is in wireless networks. One such network is shown in Figure 3 of the accompanying drawings. The broadband radio access network 1 (BRAN) is composed of a radio access network 2 having mobile terminals 4, access points 51, 51' and an air interface between, plus a mobility enhanced IP router 52(M-Router). The BRAN is connected to the core IP network which comprises the internet backbone 21 and home agents 22.

The radio access network 2 (RAN) implements all the radio dependent functionality such as radio resource management, setup and release of wireless flows, handovers and packet compression. It contains mobile terminals and access points. The mobile terminal 4 is the user's communication device for accessing wireless Internet services, and is the end point of the Internet and radio access network control protocols. The access point 51,51' implements all the radio dependent control functionality, such as radio resource management. It includes radio resource management and radio link control functions. The corresponding network elements in GSM are the base transceiver stations (BTS/TRX) and base station controller (BSC).

The M-Router 52 creates the wireless IP sub-network managing one or more access points. The M-Router handles the mobility and location management of the terminals that are registered to the access points 51,51'. The M-Router provides IP mobility services, such as DHCP (dynamic host configuration protocol). DHCP is used for allocating IP addresses for the terminals. The corresponding element to the M-router in the GSM network is the gateway GPRS support node (GGSN). The access points 51, 51' and the terminals 4 with an IP stack that belong to the same IP sub-network (use the same M-ROUTER) create a logical link.

The core network 3 comprises a home agent 22 which resides in the home network of an associated terminal 4 and is accessed through standard IP gateways. Typically home agent 22 is implemented as part of the M-Router 52 of the home network. However, it can also be a separate entity (e.g. PC host). The home agent 22 can contain user authentication information and a billing database. It resembles the home location register (HLR) in GSM.

In a preferred embodiment the M-router 52 provides IP flow classification. The network can assign certain quality of service characteristics for a flow, which are required for multimedia service implementations in IP networks. For instance, a particular flow can be prioritised in the router. In the present embodiment, the M-router maintains IP flow QoS characteristics in the air interface and permits the prioritisation of different IP packet (flows) in the radio link. It does this by mapping the detected IP flows into corresponding radio flows for transmission over the RAN. These radio flows have corresponding identifiers and QoS characteristics, and are further discussed in Finnish patent application number 980191, a copy of which is attached as Annex 1.

Alternatively, the IP flow classification could be positioned into the access point controller or even into each access point.

Moreover the criteria used for flow classification in the preferred embodiment to detect an IP flow is not essential to the invention. If IP flow detection is required, then various other criteria can be used. For example, the flow classifier can be dynamically configured by changing the value of the packets/sec detection criteria parameter.

The present invention includes any novel feature or combination of features disclosed herein either explicitly or any generalisation thereof irrespective of whether or not it relates to the claimed invention or mitigates any or all of the problems addressed.

In view of the foregoing description it would be evident to a person skilled in the art that various modifications may be made within the scope of the invention. For example, QoS characteristics (such as priority/real time/accuracy) were given as the packet management criteria in the preferred embodiment. However, packet management criteria can be based on other factors for which the network requires differentiation.

ANNEX 1.

Method supporting the quality of service of data transmission

The present invention relates to a method as set forth in the preamble of the appended claim 1 for supporting the quality of service of data transmission in wireless communication according to the Internet protocol, a system as set forth in the preamble of the appended claim 8, and a wireless communication device as set forth in the preamble of the appended claim 14.

The International Standardisation Organisation ISO has developed an open system interconnection (OSI) model for describing the distribution of data transmission in different layers. The layers are, listed from top downwards, an application layer, a presentation layer, a session layer, a transport layer, a network layer, a data link layer, and a physical layer. In view of the present specification, the most essential layers are the physical layer, the data link layer and the application layer.

The European Telecommunication Standards Institute ETSI has defined a standard for a wireless local area network (ETS 300 652), HIPERLAN Type 1 (high performance radio local area network) to be applied *e.g.* in wireless local area networks of short distances, such as local area networks of offices. In a local area network according to this standard, several devices may be connected which communicate on the same data transmission channel using packet data transmission. The standard defines the two lowermost layers of said OSI model: the physical layer and the data link layer.

The Conference of European Post and Telephone Administrations CEPT has defined a standard TR 22-05 where the frequency range from 5.15 GHz to 5.3 GHz is reserved for data transmission according to the HIPERLAN standard. This frequency range is divided into five channels, each of which being allotted a band width of ca. 23.5 MHz. Figure 1a shows a reduced example of such a local area network according to the HIPERLAN standard. It consists of terminal nodes 101a, 101b, 101c, 101d, a switching node 102 and a gateway node 103. The terminal nodes 101a—101d may communicate directly with each other, or they may communicate via the switching node 102 if there is no direct radio communication between the terminal nodes 101a—101d due to *e.g.* too long a distance or obstacles dampening radio signals. Via the switching node 102, the terminal nodes 101a—101d can also communicate with the gateway node 103 which is coupled to *e.g.* a wireless local area network 104 or the Internet network. Thus, the terminal node 101a—101d can be used as an Internet host, if necessary.

Figure 1b shows the structure of a data transfer packet according to the HIPERLAN standard. First, there is a header which is transmitted at a lower bit rate (LBR) than the other blocks and which includes the address information and the length of the packet. This is followed by a synchronisation block for synchronising the receiver to the data blocks of the packet DB(1), DB(2), ..., DB(m) containing the actual information to be transmitted. One packet may contain a maximum of 47 data blocks. Each packet can be addressed to either one receiver (unicast packet) or several receivers (multicast packet). As the third packet type the HIPERLAN standard defines an acknowledgement packet (ACK) by which the receiver of the packet informs about the successful receipt of the packet so that the sender will know if there is a need to retransmit the packet. In packets requiring data transmission in real time, it can be defined that the receipt of the packet is not acknowledged, because the information contained in the packet could be outdated if retransmitted. Packets of this kind are, for instance, packets for audio applications. On the other hand, for some real-time applications with higher quality demands, such as video applications, it is possible to define limited packet acknowledgement, whereby the acknowledgement is transmitted for several packets with one message. In packets not requiring real time, it is possible to define the acknowledgement to be sent after the receipt of each packet.

The transmission and receipt take place on the same channel without external synchronisation. The channel is listened to by the receiver of the transmitting node for a certain time, and if no communication is detected on this channel within this time, it is assumed that the channel is free and transmission is started. However, if communication is detected on this channel, the receiver is synchronised with this transmission. After the transmission, a possible acknowledgement message is waited for, and after this, an attempt for obtaining the channel can be started.

However, there may be several nodes waiting for transmission turns, whereby it may occur that several terminal devices try to transmit simultaneously. This can be solved *e.g.* so that the nodes are allotted different priorities, whereby a node with a lower priority will wait a longer time after the end of a transmission before it starts to transmit, if no communication is detected on the channel within this time.

The term "Internet" is commonly used to describe an information resource from which information can be retrieved from a data processor, such as a personal computer (PC). The data processor communicates via a modem with a telecommunication network. This information resource is distributed world-wide, comprising several storage locations which also communicate with the telecommunication network. The Internet is made operable by defining certain data communication standards and protocols, such as TCP (transfer control protocol), UPD (user datagram protocol), and IP (Internet protocol), which are used for controlling data transmission between numerous parts of the Internet. The TCP and the UDP are involved with preventing and correcting data transmission errors in the data transmitted in the Internet; the IP is involved with data structure and routing. The currently used versions of the Internet protocol are IPv4 and IPv6.

Thanks to the growing popularity of open data systems, the Transmission Control Protocol/Internet Protocol (TCP/IP) communication protocol has become a generally used protocol whereby computers of different sizes and brands can communicate with each other. TCP/IP support is currently available for almost all operating systems. The network layer protocol of TCP/IP, the Internet Protocol IP, is intended to be routed by gateways, *i.e.* routers. The routing is conducted by means of IP addresses of four bytes and routing tables. Thanks to the Internet protocol, computers using the TCP/IP can transfer messages in the routing network even to the other side of the world.

The Internet, which covers well particularly the industrialised countries, is a huge network of routers using the TCP/IP communication protocol. The largest group of users of the Internet, which was originally in scientific use only, is now firms which buy their services from commercial connection providers. In the Internet, each device has its own individual

IP address. In the Internet protocol version IPv4, the IP address consists of 32 bits, *i.e.* it is a digit of four-bytes which is divided in two parts: an organisation-specific network address and a network-specific device address. For facilitating the processing of addresses, a decimal dot notation system has been introduced, in which the addresses are indicated by digits of 8 bits separated by dots (an octet). One octet is a number from 0 to 255. This address mechanism is further divided into three different classes (ABC) which make network and device addresses of different lengths possible.

Further, with the growing popularity of the Internet, the length of the address blocks in the data packets of Internet messages is no longer sufficient in all situations for indicating all the addresses in use. This is one reason for developing the Internet protocol version IPv6. In this protocol version, the length of the address blocks is increased to 128 bits, which means in practice that an individual address can be reserved for all devices that are connected with the .Internet network. Figure 2 shows the blocks of the data packet in Internet messages.

The header block consists of the following elements:

Version	IP version of 4 bits (=6)
Prio.	4 bit priority,
Flow label	24 bit label for identifying the connection in the application layer,
Payload length	16 bit integer indicating the length of the payload, *i.e.* the length of the packet after the header in bytes,
Next header	data of 8 bits determining the header immediately following the IPv6 header,
Top limit	integer counter of 8 bits which is reduced by one at the each device (node) which transmits the packet further; the packet is rejected if the value is reduced to zero,
Source address	the 128 bit address of the sender of the original packet,
Destination address	the 128 bit address of the intended recipient.

The header is followed by the payload block, *i.e.* the actual information to be transmitted.

Physically, the Internet consists of communication network arranged in a hierarchy, for example local area networks (LAN), regional telecommunication networks, and international telecommunication networks. These communication networks are coupled internally and externally with routers which transmit information from the transmitting terminal equipment or from the preceding router in the chain of data transmission, and route the information to the receiving terminal equipment or to the next router in the chain of data transmission.

Figure 3 shows the coupling of the transmitting terminal equipment (source host, SH) and the receiving terminal equipment (destination host, DH) to the Internet via corresponding local area networks LAN and routers R.

Below in this specification, the transmitting terminal equipment and receiving terminal equipment will also be called by the common term Internet host. The Internet host can be typically used either as the source host SH and the destination host DH.

An Internet host, coupled to the Internet network via a local area network LAN, is either provided with a permanently defined Internet address or the address is a dynamic address generated by the server of the local area network (for example by using a dynamic host configuration protocol DHCP). In case the Internet host is coupled by a modem to a telecommunication network, the telecommunication terminal must ask for an Internet address from an Internet service provider to which the Internet host is registered. This is conducted *e.g.* according to a point-to-point protocol (PPP) formed above the Internet protocol layer. In both cases, the information to be transmitted in the Internet is routed to the Internet host possibly via several communication networks and routers from a remote host by using a determined Internet address.

The IP defines the transmission of the communication in packets (datagrams). The packet data transmission is one reason for the popularity of the Internet, because it allows transmission in bursts which does not require constant on-line connection and makes it possible that several Internet hosts are coupled in the same telephone connection. When a router receives a packet containing a destination address, the router routes the packet forward, if there is free capacity in the buffer memory of the router and at least one open telephone line. If there is not sufficient memory space or no open telephone line available at the moment, the packet is rejected and the source host or the preceding router must try retransmission later. In general, the Internet does not support time-critical data transmission, and the method of best effort offered by the Internet protocol is sufficient.

In the transmission of packets according to the Internet protocol, the packets can be transmitted directly to the receiver only when the network elements of the addresses of both the host and the destination are the same. In other cases, the packets are transmitted to a router which takes care of transmitting the packets further, either to the next router or to the destination, if the recipient is in the network of the router. In each router, each packet entering the router is transferred from the communication layer according to the OSI model to the network layer, where the header of the packets is examined, and on the basis of the address data therein, a decision is made where the packet is to be transmitted. For transmission, the packets are transferred back to packets of the communication layer. Because the Internet protocol has the character of a connectionless protocol, the above-mentioned operations must be taken for each packet entering the router. If the communication layer is fast, for example in accordance with the asynchronous transfer mode ATM, the processing of the packets takes a significant part of the time used for transmission. Thus, the whole transmission capacity of the transfer line cannot be utilised effectively. For correcting this situation, *e.g.* Ipsilon Networks has developed a coupling solution. In this solution, an attempt is made to detect time-consuming data transmission flows and to couple them directly with a communication layer.

The coupling solution by Ipsilon Networks consists of switches and controllers for controlling their operation. When a continuous communication flow is detected by the controller in any protocol communication in the Internet, the controller requests the transmitter to label the packets of said communication flow with a flow label, *i.e.* to open a so-called virtual channel for this communication flow. If the same finding is made by the receiver, also it requests for separation of the communication flow onto a separate virtual channel. Subsequently, this controller between the transmitter and the receiver may locally control their own switch to turn on direct communication between these two virtual channels. Because the presented coupling solution is based on labelling communication flows, it contains for each label a defined time limit after which the label is rejected, if there is no longer communication on the channel labelled by it. This reduces the number of different labels required simultaneously. In this solution, the coupling is made on the basis of communication between three nodes, and the switching request is made by the sender and/or the receiver. The coupling reduces primarily the delay of data transmission in comparison with routing.

This coupling solution is only intended for accelerating routing of packets according to the Internet protocol, and this coupling solution requires that three nodes are involved. This solution does not consider the quality of service as such.

Data transmission in packet form improves the degree of capacity utilisation of the communication channel in general, not only for retrieving information from the Internet. For example, packet data transmission can be used in applications, such as voice calls, video negotiations and other communications according to different standards. However, some of these applications are time-critical. For example in a real-time voice call, the service of best effort offered by the Internet protocol may cause significant delays in the transmission and transfer of the audio signal, which affects the understanding of the received audio signal so that *e.g.* speech is almost or totally intelligible. Moreover, the delay (the time consumed from the transmission to the receipt of the packet) may vary during the transmission of the audio signal, depending on *e.g.* the load of the communication network and variations in transmission errors. The same applies also to the transmission of a video signal in real time. There may also be situations where the users of Internet do not want as long delays as occur in many cases for obtaining information from the Internet.

The Internet Engineering Task Force (IETF) is an organisation involved with the development of Internet architecture and operation in the Internet. The IETF is currently developing a new protocol which provides an Internet host the possibility to request a desired quality of service from available defined qualities of service (QoS). This protocol is known as the resource reservation protocol (RSVP), and it is presented in the standard proposition "Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification" by Braden, R.; Zhang, L.; Berson, S.; Herzog, S.; Jamin, S.; RFC 2205, September 1997 (available at http://www.isi.edu/div7/rsvp/pub.html). The Internet host uses the RSVP protocol when requesting a certain quality of service QoS from the Internet network on the basis of the communication flow of the application which the Internet host wishes to receive from a remote host. The RSVP protocol transmits the request through the network by using each router user by the network for transmitting the communication flow to the receiving Internet host. In each router, the RSVP protocol tries to make a resource reservation for said communication flow. Also, the RSVP protocol tries to make a resource reservation for the communication flow in the destination and source Internet host.

For making a resource reservation in any node, which may be either a router or an Internet host, the RSVP protocol communicates with two local terminal modules: access confirmation module and protocol module. The access confirmation module deduces whether the node has sufficient resources for providing the requested quality of service. The protocol module deduces if the user has access to make a reservation. If either checking fails, the RSVP protocol returns an error message to the application that formed the request. If both tests are successful, the RSVP protocol sets the parameters for classifying the packet and for scheduling the packet in the Internet source host for achieving the desired quality of service. The classification of the packet deduces for all packets a service quality class, and the scheduling controls the transmission of the packets for achieving the promised quality of service in all communication flows.

The RSVP protocol operates on top of the Internet protocol both in IPv4 and IPv6. In particular, the RSVP protocol is designed to utilise the strong points of the routing algorithms of the present Internet. The

RSVP itself does not conduct routing but it uses routing protocols of lower levels to deduce where reservation requests should be transferred. Because the routing changes the routes for complying with changes in the topology of the Internet network, the RSVP protocol places its reservations for resources in new routes, if necessary.

Telecommunication networks and the Internet are two significant worldwide communication networks, whereby wireless telecommunication terminals are developed for coupling therewith and for their use. For example, cellular networks make it possible to couple a wireless telecommunication terminal to a telecommunication network and offer a high quality of service with circuit-switched technology. These cellular networks and other mobile communication networks can be utilised also for coupling to the Internet network and for utilising multimedia services. However, the circuit-switched system has the disadvantages that the connection from a wireless telecommunication terminal to a wireless communication network is turned on during the whole connection, which takes up the capacity of the wireless communication network and limits the number of simultaneous connections.

In solutions of prior art for wireless packet communications, obtaining a quality of service is not supported. Because of this, a development in the Internet community has been started for solutions supporting the mobility of Internet host and obtaining quality of service in the Internet protocol version IPv6.

In radio links, data is typically transmitted in a channel which is a certain frequency range. In one system, several channels can be available simultaneously. Further, in full duplex data transmission there are separate transmitting and receiving channels, whereby for example a base station transmits on the transmitting channel to the terminal device and the terminal device transmits on the receiving channel to the base station. A problem with radio links is that the radio channel is a limited resource which limits *e.g.* the band width and/or number of channels that can be reserved as well as the data transmission rate available for the radio link. The radio channel is liable to disturbances, such as distortion of the received signal caused by multi-channel propagation which is due to the fact that the same signal is received at the destination through different routes at different times. To reduce the effect of disturbances, part of the data transmission capacity must be used for transmitting error correction data with the packets, and achieving a desired error probability rate may require several packet retransmissions, which reduces the capacity of the radio link.

In radio links where several data transmission flows are transmitted on one channel, packets of these different data transmission flows are multiplexed. The transmission order can be affected by arranging packets of different data transmission flows in an order of priority, whereby packets of a flow with higher priority are transmitted more often than packets of a flow with lower priority. These include packets of a real-time application which are preferably made as short as possible. On the other hand, packets of applications with lower priority are often considerably longer than packets with higher priority. In systems of prior art, such a long packet prevents the transmission of other packets as long as the transmission of the packet takes. This may cause considerable delays also in the transmission of packets with higher priority, and reduce the quality of service.

It is an aim of the present invention to provide a method for flexible determination of the quality of service in wireless communication in the Internet. The method of the invention is primarily characterised in what will be presented in the characterising part of the appended claim 1. The system of the invention is primarily characterised in what will be presented in the characterising part of the appended claim 8. Further, the wireless communication device of the present invention is primarily characterised in what will be presented in the characterising part of the appended claim 14. The invention is based on the idea that for setting up an Internet connection, the required quality of service is determined for the connection, on the basis of which the connection is attempted to make in a wireless communication network with parameters complying with the set quality of service.

The present invention gives significant advantages to the solutions of prior art. In a wireless connection set up by the method of the invention, the quality of service is obtained in a more reliable way, and moreover, the whole capacity of the wireless communication network can be utilised more efficiently, because for some connections it will suffice to have a quality of service which takes up less of the capacity of the communication network. On the other hand, fewer retransmissions will be required in connections where no high demands are set for the correctness of the data transmission, *e.g.* for the transmission of speech or video between the Internet network and a wireless telecommunication terminal. Thus, more capacity will be left for applications where *e.g.* the correctness of data transmission is important, such as in the transmission of data files. Data transmission flows are examined in a centralised manner on the Internet protocol level, and a detected data transmission flow is supplemented with the parameters of the quality of service of the radio interface. These parameters are obtained advantageously from a predetermined data file depending on the radio interface. In this centralised definition of quality of service, only two parties are required, and it is possible to better examine active data transmission flows and quality of service levels defined for them, before setting the quality of service for a new data transmission flow. Thus, new connections will not reduce the quality of service of existing connections.

As another advantage, it can be also mentioned that the packet of one data transmission flow does not need to be transmitted as one packet but it can be divided into smaller parts which are, according to the invention, equipped with a label of the radio flow, on the basis of which the receiver can distinguish between packets of different flows and their parts. Thus, between parts of one long packet, it is possible to transmit a packet of a flow requiring higher quality of service. Further, the number of retransmissions can be reduced, because errors occur typically in bursts, whereby not all parts of a long packet are not necessarily erroneous and these do not need to be retransmitted.

In the following, the invention will be described in more detail with reference to the appended drawings, in which

Fig. 1a shows an example of a local area network complying with the HIPERLAN standard in a reduced manner,

Fig. 1b illustrates the structure of a data transmission packet complying with the HIPERLAN standard, Fig. 2 illustrates the structure of a packet of the Internet protocol version IPv6,

Fig. 3 shows the coupling of a transmitting and receiving host via the Internet network in a reduced chart,

Fig. 4 shows the coupling of a wireless Internet host to the Internet network in a reduced chart,

Fig. 5 shows packet data transmission between a wireless communication device and the Internet network via the GSM cellular network,

Fig. 6 shows an example of generating a radio flow label upon detecting a data transmission flow in a wireless communication network, and

Figs. 7a and 7b show examples of packet transmission sequences according to prior art and upon transmission with a radio flow label according to the invention.

In the following, the invention will be described by using the GSM cellular network as an example of a wireless communication network and a wireless communication device of the GSM system as the wireless Internet host, but the invention can also be applied in other wireless communication networks and wireless telecommunication terminals with the option for data transmission in packets. This wireless communication device 1 can also consist of a computer, such as a portable computer, coupled with a wireless data transmission device, such as a radio modem.

In this specification, data flow refers to the transmission of data packets belonging to the same communication/application. Respectively, wireless data flow refers to the transmission of data packets belonging to the same communication/application, advantageously via the radio channel, whereby also the term radio flow is used. The packets may be *e.g.* packets complying with the Internet protocol or GPRS packets of the GSM cellular network. The GPRS packet transmission system provides the possibility of 14 simultaneous connections in one terminal (wireless communication device) at the data transmission level. At present, the GPRS packet transmission offers the possibility of arranging the packets in four different levels of priority. The block reserved for this priority information in the packet can be modified into a block reserved for the radio flow label according to this invention. In case there is a need to form at least as many radio flows as the number of simultaneous connections, the corresponding number of bits are reserved for the radio flow label. Thus, two additional bits will be needed in addition to the priority block.

Each connection may be connected with one application, but the same application may involve also more than one connection. The data transmission flows of these different connections belonging to the same application can be identified by the address and port data of the sender and the receiver in the header of the packets.

Figure 4 is a reduced chart showing the coupling of a wireless Internet host in the Internet network. The system consists of a wireless communication device 1, a radio access network 2 and a core network 3. The radio access network comprises the operations for accomplishing data transmission between the wireless communication device 1 and the core network 3 as well as for controlling wireless resources, for setting up and down wireless data flows or radio flows, for moving the connection from one control station to another (handover), and possibly also for compressing packets *e.g.* according to the IPv6 standard. In this example, the functional elements of the radio access network include an access point 4, 4' (AP) and an access point controller 5 (APC). A radio communication is set up between the access point 4 and the wireless communication device 1, for transmitting *e.g.* signals required for setting up the connection and information during the connection, such as data packets of an Internet application. The access point controller 5 controls over one or several access points 4, 4' and connections set up through them to wireless communication devices 1. The radio access network 2 may comprise several access point controllers 5, 5', 5". In the GSM cellular network, the access point 4, 4' is a base station and the access point controller 5, 5', 5" a base station controller.

The core network consists of nodes connected by wires in the Internet, such as routers and wired Internet hosts.

The core network can be divided into so-called domains. These domains have a server computer or a corresponding router, by means of which the domain can communicate with other domains in the Internet. The Internet hosts in the domain, in turn, are coupled with the router of the domain. Figure 4 shows a core network with two such domains 6, 6' which are intended for serving wireless communication devices 1. These domains 6, 6' include mobile domain (MD) routers 7, 7' which control the access point controllers 5, 5', 5" coupled with the domain 6, 6'. Mobility is achieved in Internet protocol version 6 by supplementing the protocol with a data transmission method whereby the domains can transmit information from a wireless Internet host that has changed its domain. This data transmission method is called in this specification a home agent. In this context, reference is made to the Internet protocol standard version 6 IPv6, where operation of this home agent is described in more detail. The mobile domain router 7, 7' contains the functional properties of the dynamic host configuration protocol version 6 DHCPv6 and the monitoring of the mobility of the wireless communication device 1 between the access point controllers 5, 5', 5" coupled within the mobile domain 6, 6'. It should be mentioned that in some domains, there may be one or several conventional routers between the mobile domain router 7, 7' and the access point controller 5, 5', 5", even though these possible routers are not shown in the appended Fig. 4. In the GSM cellular network, where the general packet radio service GPRS is used, the element corresponding to the mobile domain router 7, 7' is the serving GPRS support node SGSN. The element corresponding to the home agent in said GSM cellular network is the gateway GPRS support node GGSN.

The network architecture used as an example in this specification gives an outline on how the quality of service can be defined in band-limited radio access networks when coupled with the Internet network. This architecture involves two connection interfaces: the radio interface and the radio access network / core network interface. Thus, the radio interface is generated for communication between the wireless communication device 1 and the access point 4, 4'. In a corresponding manner, the radio access network / core network interface consists of the connection between access point controllers 5, 5', 5" and mobile domain routers 7, 7'.

The user of the wireless communication device 1 can use the Internet network *e.g.* in a way that an application program, such as a browser, designed for this purpose is turned on in the wireless communication device 1. In the application program, the user of the wireless communication device sets as the destination address the address of a desired Internet server or Internet host, for example the address of the Internet server of the provider of the service with which the user of the wireless communication device has made a subscription to using Internet services. As already presented above in this specification, this Internet address can be given as a four-part octet number string or addresses in text form can be used, whereby a domain name server converts the address from text form into a numerical string according to the Internet protocol.

Figure 5 is a chart showing a situation where the wireless communication device 1 is coupled to the Internet network via a digital cellular network by using the general packet radio service GPRS. The wireless communication device 1 communicates with an access point 4 on any channel of the frequency range reserved for the system. In the GSM cellular network, this access point 4 is a base transceiver station (BTS) of the base station subsystem (BSS). One access point 4 forms the radio interface of one cell in the cellular network. The access point 4 operates as a transmitter of information to be transmitted between the wireless communication device 1 and the access point controller 5. It is a central function of the access point controller to control the channels in the interface and to transmit the connection from one access point 4 to another access point 4' in a situation when the wireless communication device 1 moves from one cell to another.

Next, data transmission from another Internet host to the wireless communication device 1 will be described. The Internet application of the wireless communication device 1, to which the information is finally transferred, transmits the above-mentioned address to define the source Internet host. The data transmission is conducted according to the GPRS standard from the mobile station 1 to the GSM cellular network. The GSM cellular network converts the packet message to a message complying with the Internet protocol and transmits it to the Internet network. The information formed in the application is transmitted to the wireless communication device 1 according to the Internet protocol via the Internet network in a manner known as such by routing to the GSM cellular network, where the information is converted to comply with the packet transmission mechanisms of the cellular network, in this case into packets of the GPRS network. The information is transmitted further via the access point controller 5 to the access point 4 and further to the wireless communication device 1 where the received message is transferred to the application layer to be processed by the application.

The following is a description on the method according to an advantageous embodiment of the invention for generating a radio flow label in communication between the wireless communication device 1 and the access point 4, 4'. The application is an Internet application of the wireless communication device 1, from which information complying with the Internet protocol is transmitted to the Internet network. This specification does not contain a more detail description on the formation of packets between the wireless communication device 1 and the mobile communication network, which may vary in different mobile communication networks and is prior art known as such by an expert in the field. Figure 6 is a schematic diagram of this formation of the radio flow label for data transmission between the wireless communication device 1 and the access point controller 5. All data transmission is based on packets and is routed according to the Internet protocol. The mobile terminal radio flow agent (MRFA), which is implemented advantageously in the application software of the wireless communication device 1, starts to transmit radio flow information packets using a default flow ID. At the access point 4, an access point radio flow agent (ARFA) transmits the flow further to the access point controller 5. At the access point controller 5, a router matrix (RM, not shown) transmits the flow to a radio flow manager block (RFM). The access point controller 5 detects that this flow is of the kind for which a radio flow label should be formed for achieving a certain quality of service (block 601 in Fig. 6). The access point controller 5 finds out if there are sufficient resources available at the moment to be used for data transmission between the wireless communication device 1 and the access point 4 in order to achieve the desired quality of service for said flow FID (block 602). If sufficient resources are available, the radio flow manager RFM selects a new flow label for the flow to be transmitted via the access point 4 to the mobile terminal radio flow agent MRFA of the wireless communication device 1. In the selection of the flow label, TCP/IP ports and/or addresses of the source host and the destination host are used. This flow label is for example data of 20 bits transmitted to the wireless communication device 1 via the access point 4. In Fig. 6, this step is indicated by arrow 603, and although it is connected directly from the access point controller 5 to the wireless communication device 1, in practical applications it is transmitted physically via the access point 4. In the wireless communication device 1, this received flow label is processed, and on the basis of this, the wireless communication device 1 generates a shorter flow label, in this application example a flow label of 8 bits, wherein a total of 256 different flow labels can be used simultaneously for different Internet applications in one wireless communication device 1.

The access point controller 5 transmits the same flow label also to the access point 4 (arrow 604); in addition, information can be transmitted here on what kind of a quality of service is desired for this flow.

The shorter flow label generated in the wireless communication device 1, which in this specification will be called the radio flow identification (RFID), is transmitted from the wireless communication device 1 via the radio interface MT/RP of the wireless communication device to the access point 4. As known, each wireless communication device of the cellular network is equipped with a device identification or a corresponding separate identification whereby wireless communication devices of the cellular system can be separated from each other. The radio interface MT/RP of the wireless communication device includes, in a manner known as such, a radio transceiver (not shown) as well as coding/decoding means (not shown), but it will not be necessary to describe this radio interface in more detail in this context. This mobile station identification MSID, which in the GSM system is advantageously the international mobile equipment identity IMEI, is transmitted from the wireless communication device 1 to the access point 4 in connection with the transmission of messages (arrows 605 and 606). Now, the access point 4 has the flow identification FID, the radio flow identification RFID as well as the mobile station identification MSID. After this, on the basis of radio flow identifications RFID coming from the wireless communication device and the mobile station identification MSID, the access point 4 can couple the flow with the original wider flow identification FID. The access point 4 transmits an acknowledgement message to the wireless communication device 1 (arrows 607 and 608) and to the access point controller 5 (arrow 609). After this, also the wireless communication device sends an acknowledgement to the access point controller 5 (arrow 610). Now, there is a connection corresponding with the desired quality of service between the wireless communication device 1 and the access point controller 5 (this is shown by block 611).

Also, the access point controller 5 may receive from the Internet network a data flow addressed to the Internet application of the wireless communication device 1. Thus, the access point controller 5 finds that a flow label can be defined for this flow, whereby the access point controller 5 examines the quality of service desired for the flow and finds out if there are sufficient resources available for achieving and maintaining the desired quality of service. At this point, the access point controller 5 considers also the other radio flows active at the moment and finds out if the desired quality of service can be provided for this flow without risking the quality of service of the active flows. If the quality of service can be achieved, the above-mentioned signalling is conducted, whereby *e.g.* a flow ID is defined for the radio flow.

In case there are no sufficient resources available on the radio channel for achieving the desired quality of service, it is possible *e.g.* to continue the radio flow at a level with a poorer quality of service, for example with a transmission of best effort, whereby the source host of the flow is informed of this procedure. If necessary, the user can be inquired if the data is to be transmitted in spite of the lower quality of service or if the data transmission is to be interrupted.

The information transmitted from the second host according to the Internet protocol is transmitted via normal mechanisms of the Internet protocol to the cellular network. In the cellular network, the message is converted to a message corresponding with the packet transmission mechanisms of the cellular network and transmitted to the access point controller 5. The access point controller 5 provides the message with a flow identification FID and transmits the message further to the access point 4. At the access point 4, it is examined on the basis of this flow identification FID what are the corresponding radio flow identification RFID and mobile station identification MSID. Next, the flow identification FID is removed at the access point 4 and replaced by the radio flow identification RFID. This way it is possible to reduce the information to be transmitted along with the packets (in this example 20 — 8 = 12 bits), which reduces the load of the radio network and makes it possible to utilise the radio network more efficiently. This is also illustrated in the appended Fig. 7a showing four transmission strings 701, 702, 703, 704 containing packets of radio flows. As examples, the packets of each string are indicated by the number of the connection (1 to 7) to which the packet belongs. Of these strings, the access point controller 5, 5', 5" selects the packet to be transmitted at each time on the basis of predetermined criteria. Prior art is shown by the first transmission sequence 705 where the order of transmission is determined primarily on the basis of priority set for the string. In this example, the order of priority is the following: the highest priority belongs to the string 701, next to the second string 702, third to the string 703, and the lowest priority to the string 704. Header blocks are indicated by letters H in each packet.

Data transmission according to an advantageous embodiment of the invention is illustrated by the second transmission sequence 706. In this situation, the transmission order of the strings 701 to 704 is determined according to the quality of service set for the radio flow corresponding to the string in a way that the higher quality of service is set for the first string 701, the next highest to the second string 702, next to the third string 703, and the lowest quality of service is set to the fourth string 704. The radio flow identifications are indicated in this second transmission sequence 706 with the reference numeral 707.

The wireless communication device 1 receives a packet message according to this transmission sequence and transmits the information contained in it to the corresponding application. The wireless communication device 1 contains also a switching table or the like containing information on the application to which a certain radio flow identification RFID corresponds. Also transmission from the wireless communication device 1 to the Internet network is conducted in a reverse order, applying the same principle.

In the formation of the packet transmission sequence, it is possible to consider *e.g.* the number of strings 701 to 704, retransmission needs caused by errors, statistical multiplexing for packets of fixed size, an attempt to reduce the average delay, and utilisation of the channel as efficiently as possible.

For defining the quality of service QoS, it is possible to utilise information in the header of the application received in the Internet message. At the present, a standard is under development on how these qualities of service could be presented and what they could be. In any case, a message according to the Internet protocol contains, in the header, information about the type of the application, which can be *e.g.* an audio application, a video application, a data application, or a combination of these. These applications of different types have different requirements. For example, the real-time processing of audio and video applications usually requires that the packets must be transmitted to the destination within a certain response time or otherwise the packets must be rejected. However, in data transmission, for example in the transmission of program files, it is the correctness, not real-time processing, of data transmission that is important. In presently known methods and cellular networks, it is defined at the design stage, what is the error probability of data transmission, on the basis of which it is possible to select error correction algorithms and to set *e.g.* a maximum number of retransmissions. All packet information is transmitted according to the same criteria. If any packet is transmitted incorrectly, it is retransmitted. These retransmissions are conducted either as long as the packet is received correctly or, if a response time is defined for the packet, the packet is rejected if it cannot be received within the prescribed time or the maximum number of retransmissions is exceeded. Since in audio and video applications even a partly incorrectly received information would be sufficient, this retransmission constitutes an unnecessary load on the radio network. On the other hand, the additional load reduces the radio resources available for other applications and thus interferes also with the quality of service obtained by other applications. For detecting and correcting errors, several methods have been developed which are prior art to an expert in the field, wherein it is rendered unnecessary to discuss them in more detail in this context. It should be further mentioned that increasing error detection and error correction capacity by error detection and correction algorithms will increase the need of data transmission. These conflicting demands set a limit to the fact how efficient an algorithm is selected, to prevent an unnecessary delay in the data transmission.

When using a method of the invention, it is possible to define different qualities of service with different demands. For example, a poorer error probability demand can be defined for audio and video packets than for data packets. On the other hand, due to the real time demand, a higher priority can be determined for audio and video packets than for data packets. Thus, data packets are transmitted at a slower rate, if the radio network is loaded. Further criteria describing the quality of service may include response time, within which the packet must be received or else it is rejected. By combining these different criteria, several different qualities of service are obtained, and also other criteria than those mentioned above can be used in defining the quality of service.

These qualities of service and the corresponding bits of the header to be examined are *e.g.* listed in a table by the access point controller 5, whereby by examining these header bits, the access point controller 5 retrieves the corresponding quality of service from the table. For these qualities of service, information is stored in the access point controller 5 on the special demands of each quality of service, including the abovementioned error probability, priority and response time.

These definitions for the quality of service are transmitted from the access point controller 5 to the access point 4 which, on the basis thereof, conducts the definition of the transmission order of the packets to be transmitted. There may be several Internet applications to be transmitted by one access point 4 simultaneously. For these different applications, a string is preferably formed for each, where packets are transferred for transmission. From these packets in different strings, the access point 4 selects the packet to be transmitted at the time.

According to the invention, it is possible to use the radio flow label to improve the efficiency of the system also in a way that the transmission of long packets can be divided into parts so that, if necessary, one or several packets of a higher quality of service are transmitted between the parts. Such a part can be *e.g.* in a time-division radio link one time period. In systems of prior art, the whole packet must be transmitted in subsequent time periods, because the receiver cannot otherwise identify the flow to which the packet part belongs. In the system of the invention, the packet parts can be identified on the basis of the radio flow identification. This situation is illustrated in the appended Fig. 7b showing four strings. Each string contains one or more packets to be transmitted. The transmission of prior art is illustrated in the first transmission sequence 705, and the transmission of packets equipped with a radio flow identification according to the invention is illustrated in the second transmission sequence 706. Thus, retransmission of so many time periods will not be needed, because, instead of retransmitting the whole packet, only the incorrectly received part or parts of the packet are retransmitted.

Determination of the quality of service according to the invention can be used also in other packet data transmission protocols and information networks. Also, in addition to the routings known from Internet networks, the invention can be applied in coupling solutions developed for Internet networks where the router is used for examining the route between data flows and conducting the coupling in the hardware layer.

The method described above as the method supporting the quality of service is applicable also together with the Internet resource reservation protocol RSVP. Thus, in the access point controller 5, 5', 5" which monitors data transmission flows, it is possible to consider also the data contained in the data transmission flow about the quality of service presented by the host. The radio flow manager block RFM formed in the access point controller 5, 5', 5" stores the parameters of the quality of service requested by the host and finds out whether the requested quality of service is available. If the requested quality of service is available, it is possible to set the parameters corresponding to the desired quality of service for the data flow in question.

The invention is not limited solely to the embodiments presented above, but it can be modified within the scope of the appended claims.

Claims:

1. A method for supporting the quality of service (QoS) in packet data transmission between a wireless communication device (1) communicating with a radio network, and an information network (LN), where data transmission between the wireless communication device (1) and the radio network (2) is controlled with at least one access point controller (5, 5', 5"), and in which method information is transmitted between the wireless communication device (1) and the access point controller (5, 5', 5") in radio flows, **characterised** in that in the method, at least one radio flow is provided with a defined radio flow identification (RFID) and a quality of service (QoS).

2. The method according to claim 1, **characterised** in that the quality of service (QoS) is determined in a centralised manner, preferably by the access point controller (5, 5', 5").

3. The method according to claim 1 or 2, **characterised** in that for determining the quality of service (QoS), the content of the packets, preferably the content of the header (H) of the packets, is used.

4. The method according to claim 3, **characterised** in that the data transmission is divided at least into a network layer and a physical layer, wherein in the method, the data transmission is conducted in packets of the network layer, which are converted into packets of the physical layer to be transmitted in a radio flow, and that the quality of service (QoS) is determined on the basis of the contents of the packets of the network layer.

5. The method according to any of the claims 1 to 4, **characterised** in that the packets of the radio flow are formed from packets complying with the Internet protocol.

6. The method according to any of the claims 1 to 5, **characterised** in that the packets of the radio flow are transmitted in the radio network (2) as GPRS packets.

7. The method according to any of the claims 1 to 6, **characterised** in that the method comprises the steps of:
— transmitting several different radio flows in packet data transmission between the wireless communication device (1) and the radio network (2), and
— transmitting a packet of a second radio flow between packets of a first radio flow.

8. A system for supporting the quality of service (QoS) in packet data transmission in a radio network (2), the system comprising:
— at least one wireless communication device (1) communicating with the radio network (2),
— means (7, 103, GGSN) for transmitting information between the radio network (2) and the information network (3, LN),
— means (5, 5', 5") for controlling data transmission between the wireless communication device (1) and the radio network (2), and
— means (4, 4', 102) for transmitting information between the wireless communication device (1) and the access point controller (5, 5', 5") in radio flows,

characterised in that the system comprises further:
— means (5, 5', 5", 103, RFM) for determining a radio flow identification (RFID) for at least one radio flow, and
— means (5, 5', 5") for determining the quality of service (QoS) for the radio flow.

9. The system according to claim 8, **characterised** in that it comprises means (5, 5', 5") for determining the quality of service (QoS) in a centralised manner.

10. The system according to claim 8 or 9, **characterised** in that it comprises means (RFM) for determining the quality of service (QoS) on the basis of the contents of the packets, preferably the contents in the header (H) of the packets.

11. The system according to claim 8, 9 or 10, **characterised** in that it comprises means (7, 103) for generating packets of a radio flow from packets complying with the Internet protocol.

12. The system according to any of the claims 8 to 11, **characterised** in that it comprises means (GGSN, SGSN) for conducting data transmission in the radio network (2) in GPRS packets.

13. The system according to any of the claims 8 to 12, **characterised** in that it comprises:

— means for transmitting at least a first and a second radio flow in packet data transmission between the wireless communication device (1) and the radio network (2), and

— means (5, 5', 5") for transmitting a packet of the second radio flow between packets of the first radio flow.

14. A wireless communication device (1) equipped with means for transmitting information into a radio network (2), comprising:

— means (7, 103, GGSN) for transmitting information between a radio network (2) and an information network (3, LN),

— means (5, 5', 5") for controlling data transmission between the wireless communication device (1) and the radio network (2), and

— means (4, 4', 102) for transmitting information between the wireless communication device (1) and the access point controller (5, 5', 5") in radio flows,

characterised in that the wireless communication device (1) comprises further:

— means (MRFA) for generating a radio flow identification (RFID) for at least one radio flow, and

— means (MRFA) for connecting said radio flow identification (RFID) into packets of said radio flow transmitted from the wireless communication device (1).

Abstract

The invention relates to a method for supporting the quality of service (QoS) in packet data transmission between a wireless communication device (1) communicating with a radio network, and an information network (LN), where data transmission between the wireless communication device (1) and the radio network (2) is controlled with at least one access point controller (5, 5', 5"). Further, in the method, information is transmitted between the wireless communication device (1) and the access point controller (5, 5', 5") in radio flows. In the method, at least one radio flow is provided with a defined radio flow identification (RFID) and a quality of service (QoS).

Fig. 1a

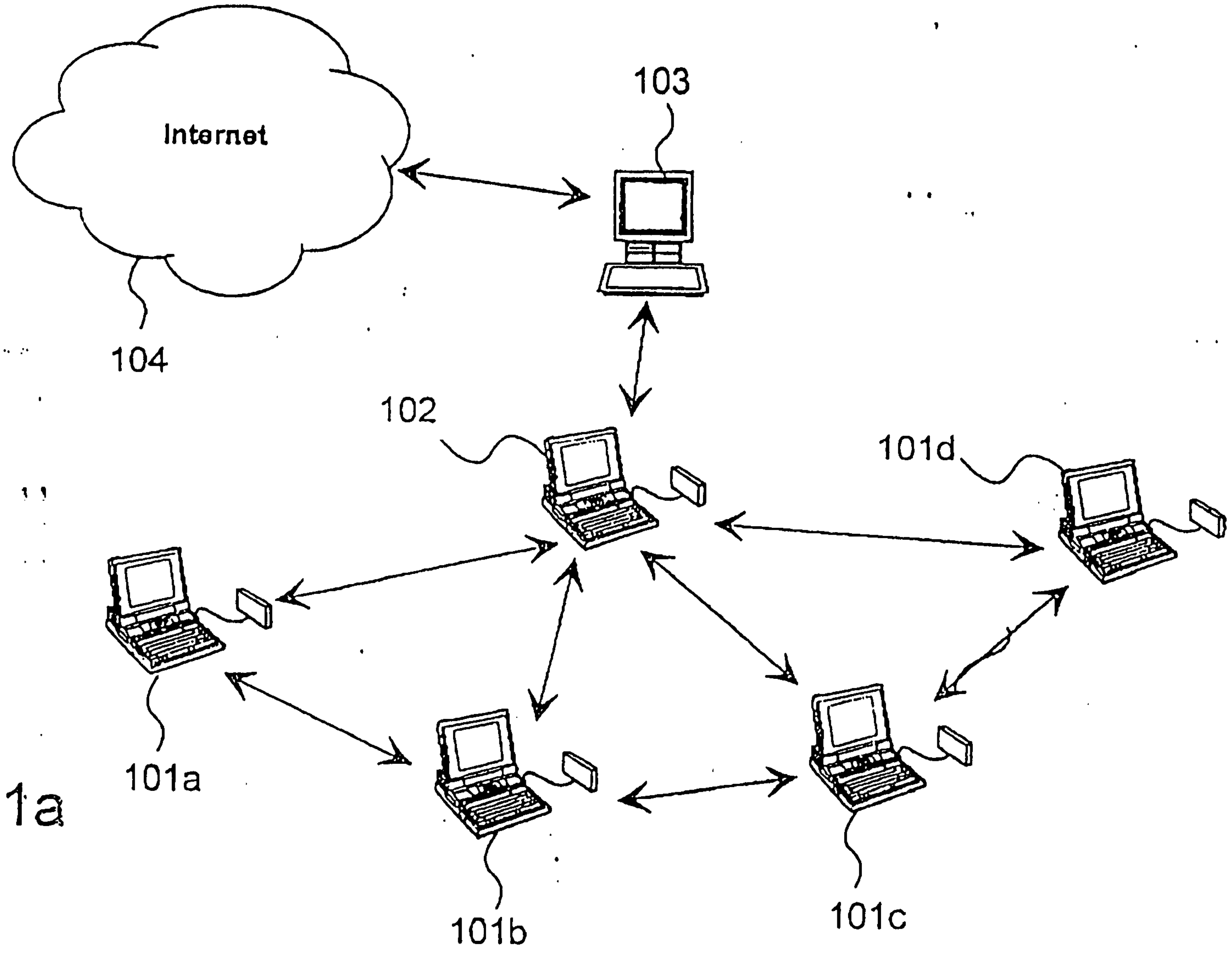


Fig. 1a

LBR Header + physical data layer data packet

8 bil Clock and Sync.	8 bil Address	2 bil Protection	6 bil Packet length	2 bil Protection	496 bil Data block 0..46

LBR Acknowledgment

8 bil Clock and Sync.	8 bil Return Address	2 bil Protection	6 bil Acknowledgment	2 bil Protection

Fig. 1b

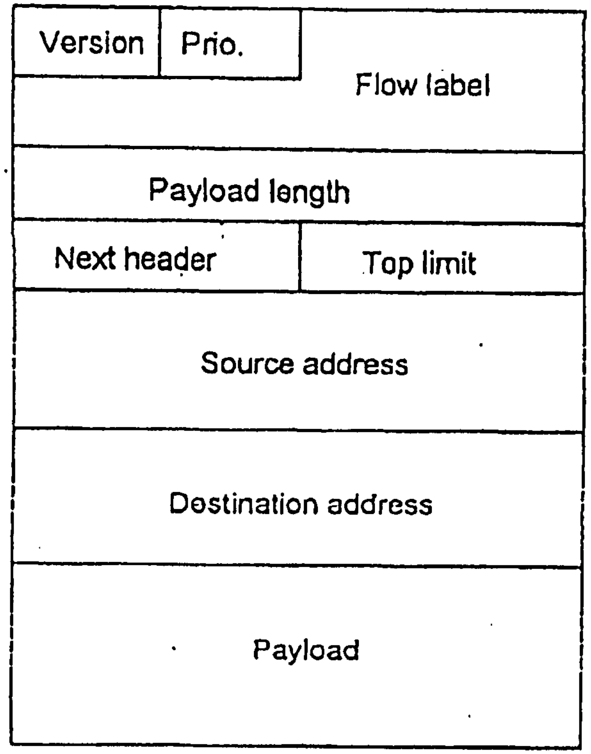

Version
Prio.
Flow label
Payload length
Next header
Top limit
Source address
Destination address
Payload

Fig. 2

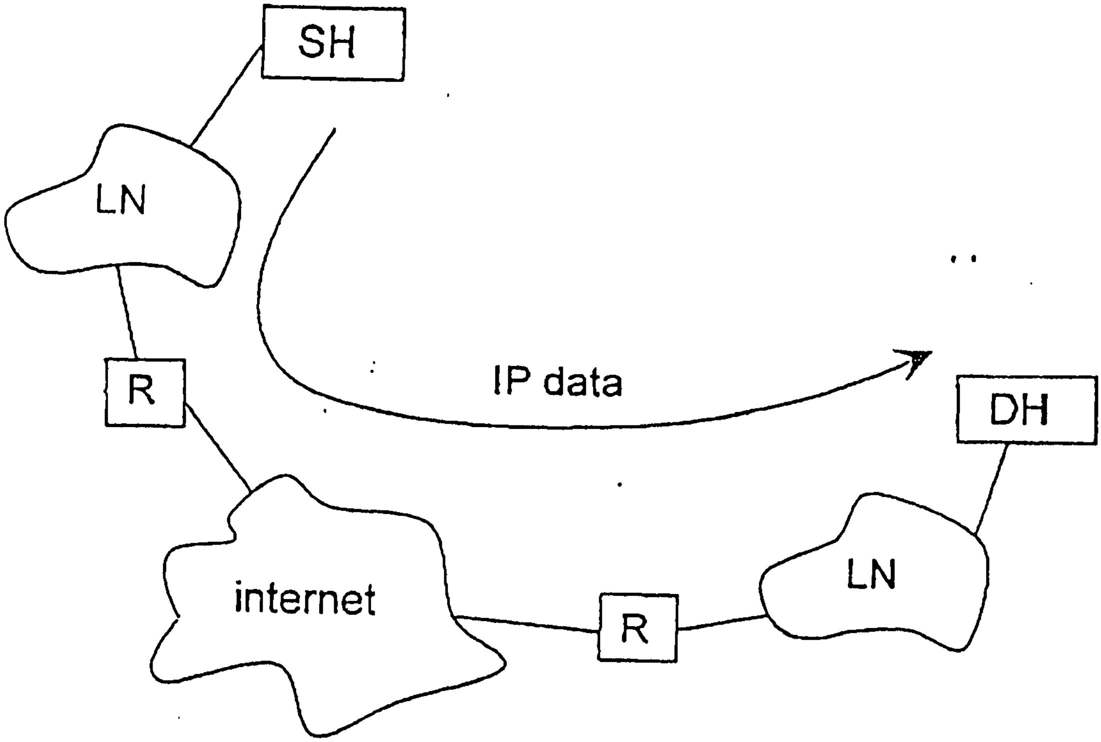


Fig. 3

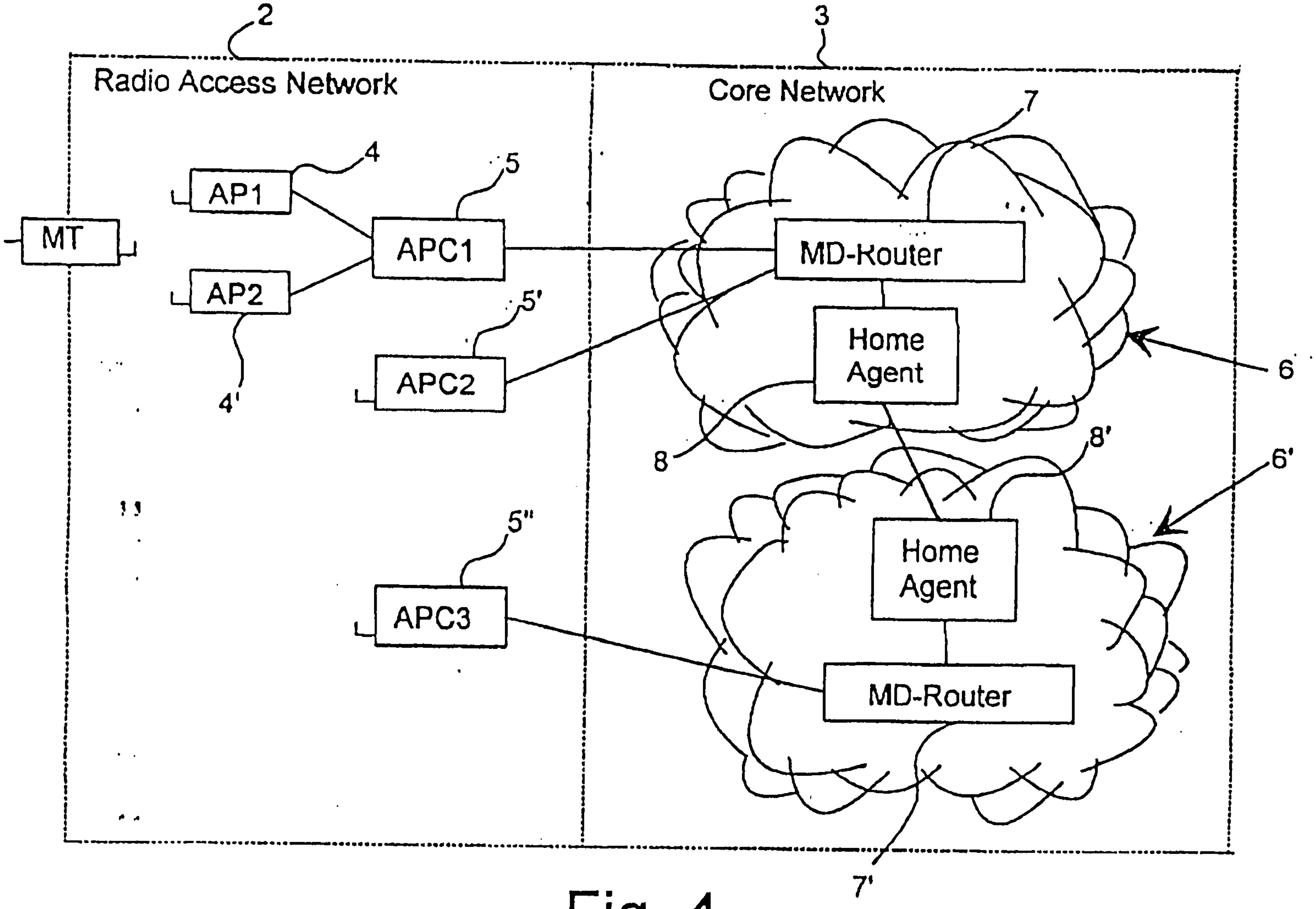

2
Radio Access Network
3
Core Network
MT
AP1
4
AP2
4'
APC1
5
APC2
5'
APC3
5"
7
MD-Router
Home
Agent
6
8
8'
6'
Home
Agent
MD-Router
7'

Fig. 4

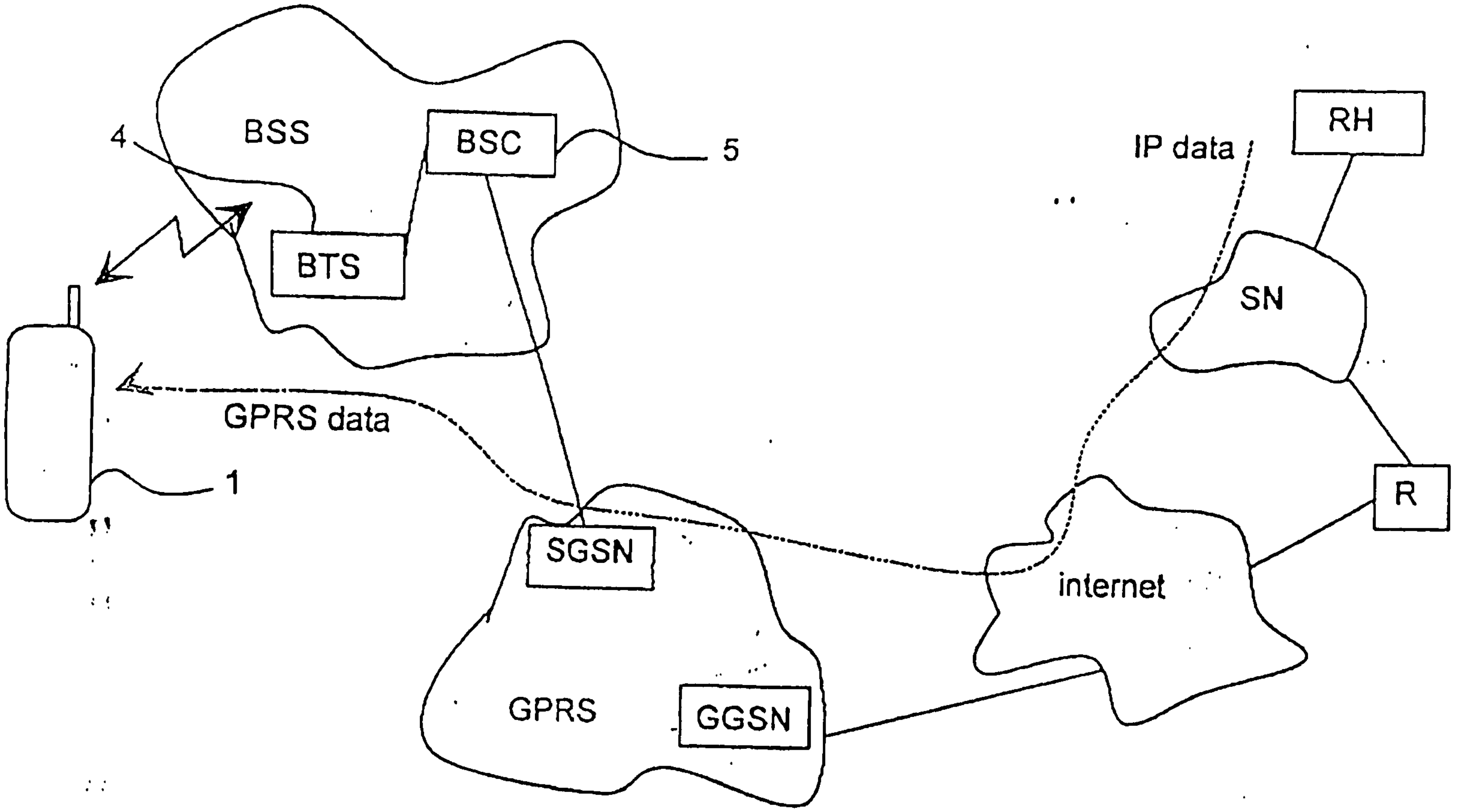

BSS
BSC
BTS
4
5
1
GPRS data
SGSN
GPRS
GGSN
internet
IP data
RH
SN
R

Fig. 5

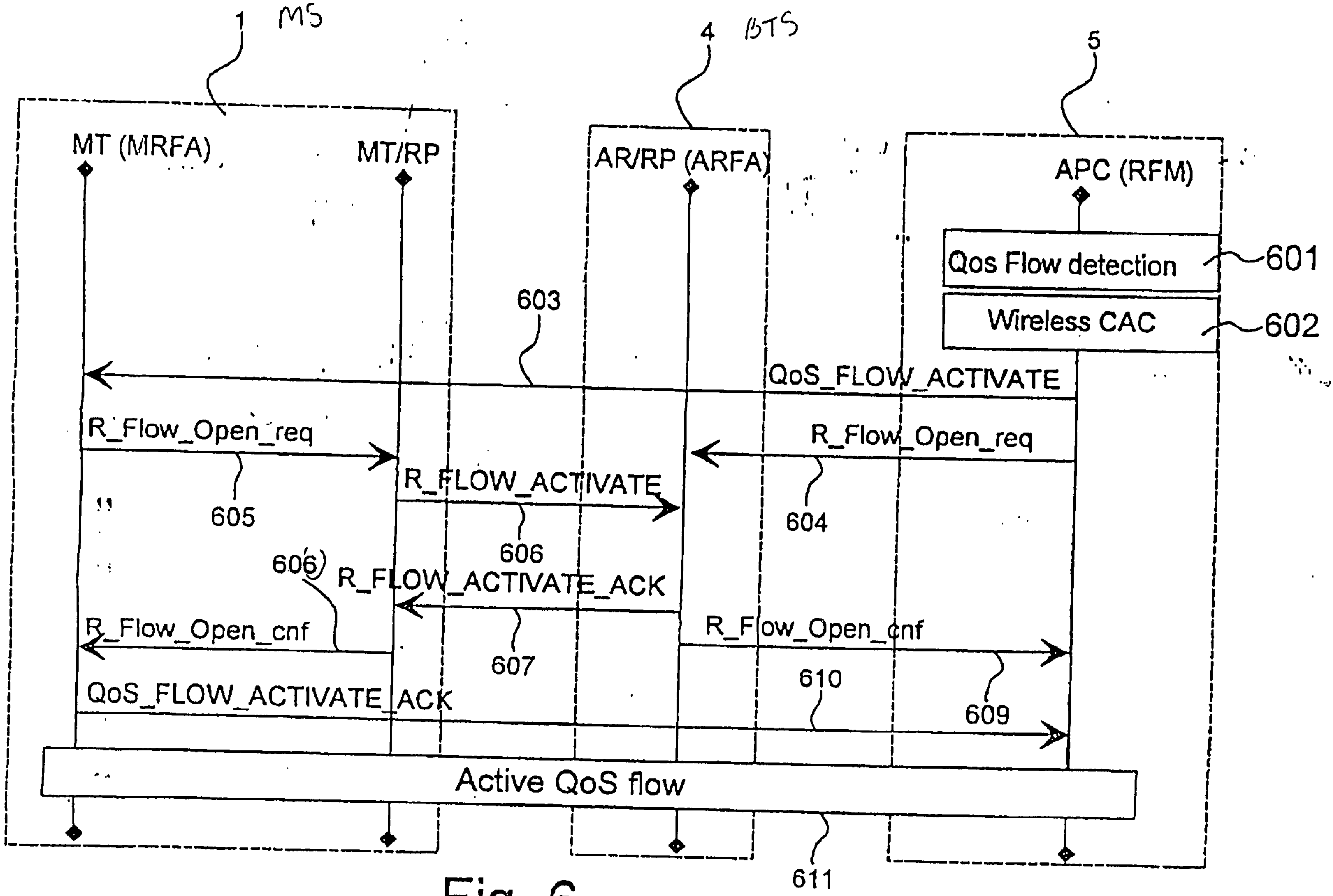

1 MS
4 BTS
5
MT (MRFA)
MT/RP
AR/RP (ARFA)
APC (RFM)
Qos Flow detection
601
Wireless CAC
602
603
QoS_FLOW_ACTIVATE
R_Flow_Open_req
605
R_Flow_Open_req
604
R_FLOW_ACTIVATE
606
606)
R_FLOW_ACTIVATE_ACK
607
R_Flow_Open_cnf
R_Flow_Open_cnf
609
610
QoS_FLOW_ACTIVATE_ACK
Active QoS flow
611

Fig. 6

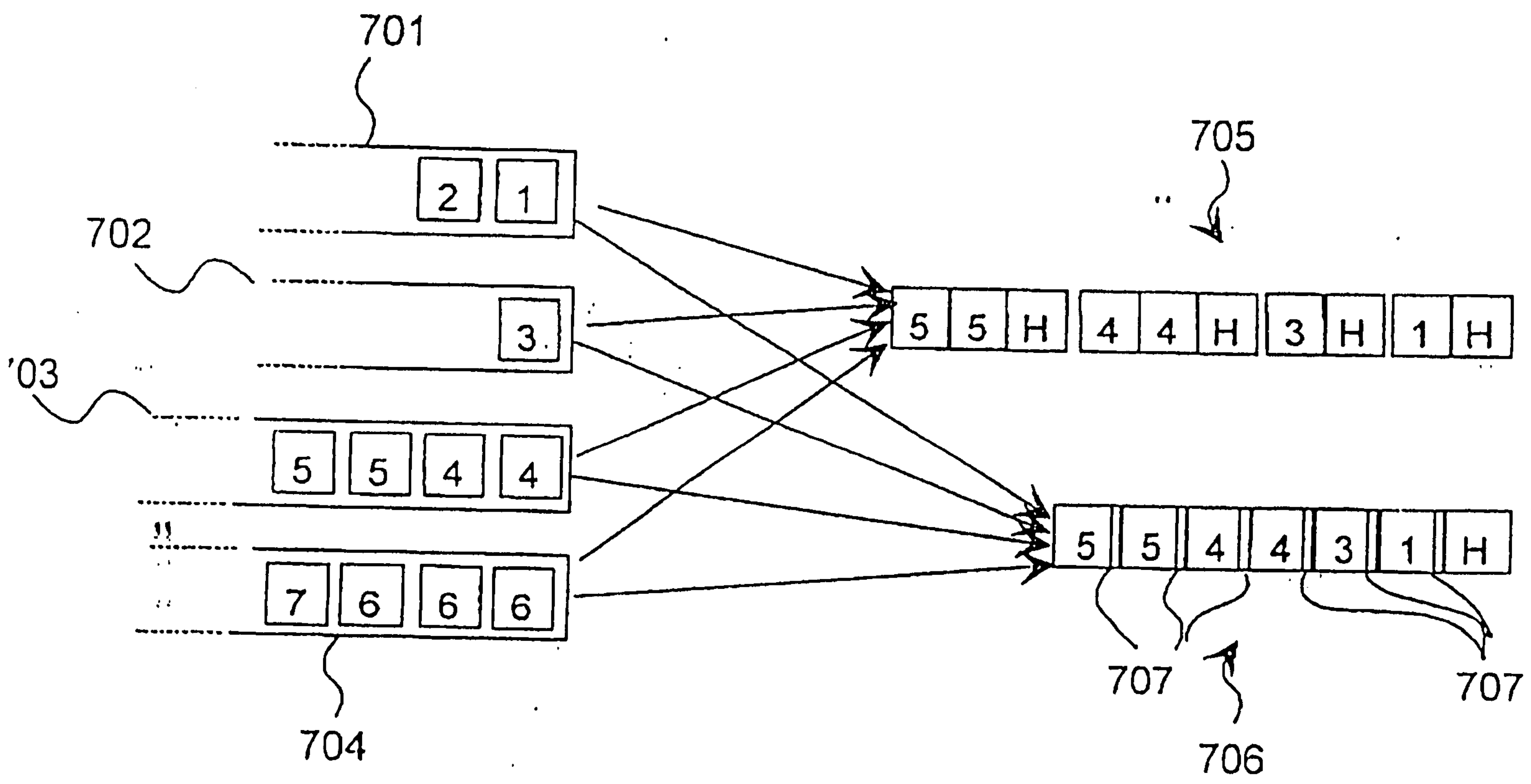

701
702
'03
704
2
1
3.
5
5
4
4
7
6
6
6
705
5
5
H
4
4
H
3
H
1
H
5
5
4
4
3
1
H
707
707
706

Fig. 7a

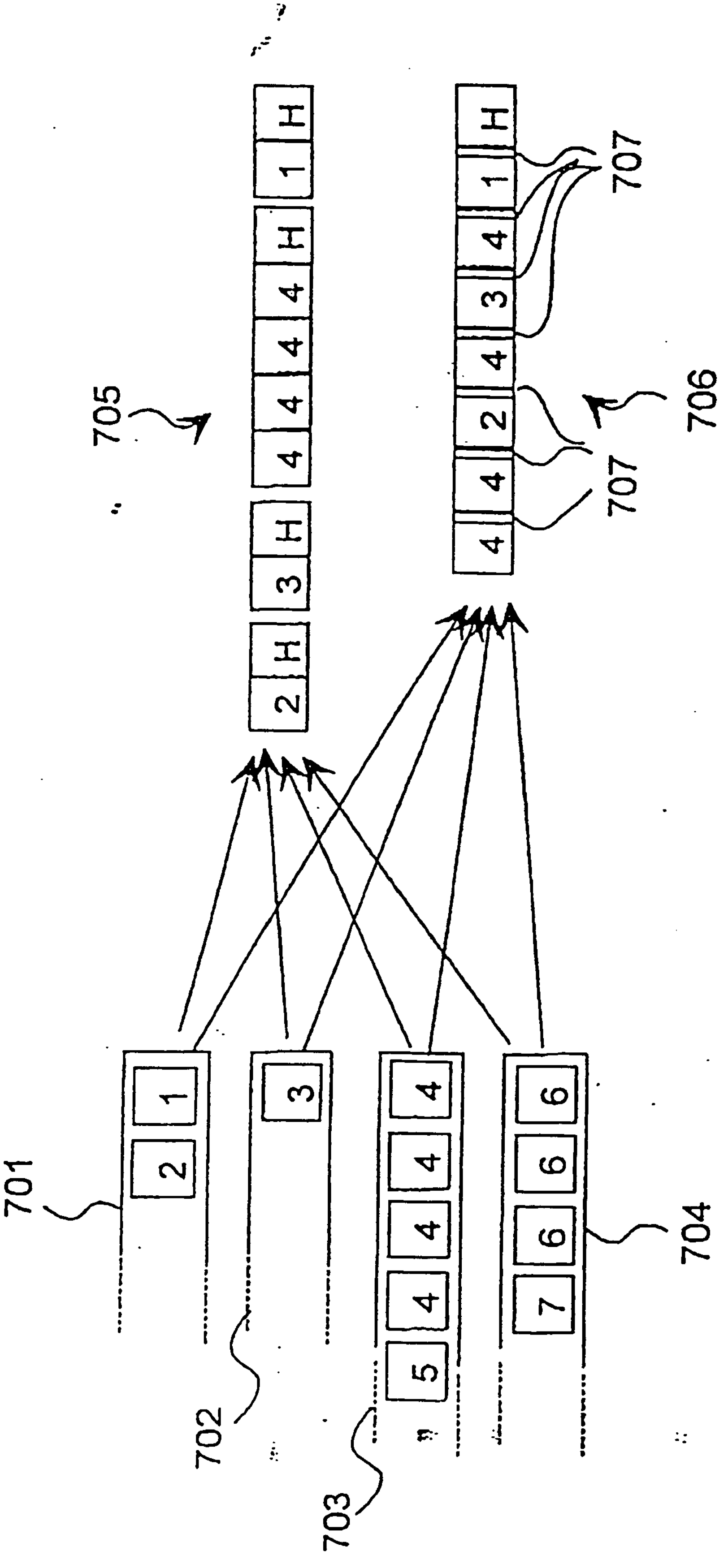

701
702
703
704
705
706
707
707
707
1
2
3
4
4
4
4
5
6
6
6
7
2
H
3
H
4
4
4
4
H
1
H
4
4
2
4
3
4
1
H

Fig. 7b

Claims

1. A method for detecting an IP flow in flow label deprived packet data transmission, comprising monitoring a set of fields in a lower layer header of the packets to detect an IP flow, wherein monitoring the set of fields comprises:

monitoring an source address field;

monitoring a destination address field; and

monitoring a further field indicative of packet management criteria.

2. A method of detecting an IP flow in packet data transmission, comprising:

selecting a set of fields to be monitored to detect an IP flow; and

monitoring the selected set of fields in a header of the packets to detect an IP flow;

wherein the set of fields is selected from:

a first set comprising a flow label field and a source address field from a lower layer header of the packets;

a second set comprising a source address field and a destination address field from a lower layer header of the packets and a source port field and a destination port field from an upper layer header of the packets; and

a third set comprising the source address field, destination address field, and a further field indicative of packet management criteria other than the flow label field from the lower layer header of the packets.

3. A method as claimed in claim 2, wherein the set of fields are selected on the basis of priority, in which the first set has the highest priority, followed by the second set, and then the third set.

4. A method as claimed in claim 2 or 3, wherein selection of the set of fields to be monitored is determined by the availability of the fields.

5. A method as claimed in any of claims 2 to 4, wherein monitoring the third set of fields comprises:

monitoring an source address field;

monitoring a destination address field; and

monitoring a further field indicative of packet management criteria other than the flow label field.

6. A method as claimed in claim 1 or 5 for detecting an IP flow in encrypted packet data transmission, wherein monitoring the further field comprises monitoring a security field.

7. A method as claimed in claim 6, wherein the encryption utilises encapsulating security payload, and monitoring the further field comprises monitoring a security field in an encapsulating security payload header.

8. A method as claimed in claim 7, wherein monitoring the further field comprises monitoring a security parameter index of the encapsulating security payload header.

9. A method as claimed in any preceding claim, wherein monitoring the set of fields comprises monitoring the set of fields in the basic and extension headers of the packets.

10. Use of a set of fields in a lower layer header of data packets as a flow identifier, wherein the set of fields comprises source and destination address

fields and a field indicative of packet management criteria other than a flow label field.

11. An IP flow detector for detecting an IP flow in flow label deprived packet data transmission, comprising a monitor for monitoring a set of fields in a lower layer header of the packets to detect an IP flow, wherein the monitoring means comprises:

a monitor for monitoring an source address field;

a monitor for monitoring a destination address field; and

a monitor for monitoring a further field indicative of packet management criteria.

12. An IP flow detector for detecting an IP flow in packet data transmission, comprising:

a selector for selecting a set of fields to be monitored to detect an IP flow; and

a monitor for monitoring the selected set of fields in a header of the packets to detect an IP flow;

wherein the selector selects the set of fields from:

a first set comprising a flow label field and a source address field from a lower layer header of the packets;

a second set comprising a source address field and a destination address field from a lower layer header of the packets and a source port field and a destination port field from an upper layer header of the packets; and

a third set comprising the source address field, destination address field, and a further field indicative of packet management criteria from the lower layer header of the packets other than the flow label field

13. A detector as claimed in claim 12, wherein selector selects the set of fields on the basis of priority, in which the first set has the highest priority, followed by the second set, and then the third set.

14. A detector as claimed in claim 12 or 13, wherein the selector determines the set of fields to be monitored on the availability of the fields.

15. A method as claimed in any of claims 12 to 14, wherein the monitor for monitoring the third set of fields comprises:

a monitor for monitoring an source address field;

a monitor for monitoring a destination address field; and

a monitor for monitoring a further field indicative of packet management criteria other than the flow label field.

16. A detector as claimed in claim 11 or 16 for detecting an IP flow in encrypted packet data transmission, wherein the monitor for monitoring the further field is arranged to monitor a security field.

17. A detector as claimed in claim 2, wherein the encryption utilises encapsulating security payload, and the monitor for monitoring the further field is arranged to monitor a security field in an encapsulating security payload header.

18. A detector as claimed in claim 3, wherein the monitor for monitoring the further field is arranged to monitor a security parameter index of the encapsulating security payload header.

19. A detector as claimed in any preceding claim, wherein the monitor for monitoring the set of fields in a lower layer header of the packets to detect an

IP flow is arranged to monitor the set of fields in the basic and extension headers of the packets.

20. A method of detecting an IP flow in packet data transmission, substantially as hereinbefore described, with reference to, and/or as illustrated in any one, or any combination of Figures 1 to 3 of the accompanying drawings.

21. An IP flow detector for detecting an IP flow in packet data transmission, substantially as hereinbefore described, with reference to, and/or as illustrated in any one, or any combination of Figures 1 to 3 of the accompanying drawings.

22. Use of a set of fields in a lower layer header as a flow detector substantially as hereinbefore described, with reference to, and/or as illustrated in any one, or any combination of Figures 1 to 3 of the accompanying drawings.

CHARTER MARK

The Patent Office

58

INVESTOR IN PEOPLE

Application No:	GB 9821994.2	**Examiner:**	Keith Williams
Claims searched:	1-22	**Date of search:**	13 April 1999

Patents Act 1977
Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.Q): H4P (PEUX, PFD, PPG, PPS)

Int Cl (Ed.6): H04L 12/24, 12/26, 12/56, 29/06

Other: Online EPODOC

Documents considered to be relevant:

Category	Identity of document and relevant passage		Relevant to claims
A	GB 2248535 A	GPT - see page 3, lines 15-19; page 8, lines 8-12	1,2,10,11 and 12
A	EP 0848527 A1	AT&T Corp. - see column 2, line 8 to column 3, line 6	1,2,10,11 and 12

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.